DOCUMENT RESUME

ED 093 909 TM 003 725

AUTHOR	Wile, Marcia Z.
TITLE	An Evolving Product Evaluation: Formative and Summative Components.
SPONS AGENCY	National Library of Medicine (DHEW), Bethesda, Md.
PUB DATE	[Apr 74]
NOTE	19p.; Paper presented at the Annual Meeting of the American Educational Research Association (59th, Chicago, Illinois, April 1974)
EDRS PRICE	MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS	Dental Schools; *Formative Evaluation; *Instructional Materials; *Medical Education; *Models; *Summative Evaluation; Textbook Evaluation

ABSTRACT

A user-centered educational product evaluation model for the prepublication and postpublication assessment of materials designed for medical and dental education is presented. The model has four components: the author, user, evaluator, and publisher. It is dynamic, reactive, and interactive, with feedback from each component providing input to regenerate the process. Three general approaches to product evaluation based on this model, for the formative and summative evaluation of printed material are described. Examples are given which illustrate the developmental nature of these educational products and their subsequent modifications, based on feedback from field tests, reviews, and experimental studies. (Author)

ED 093909

TM 003 225

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

An Evolving Product Evaluation: Formative and Summative Components

Marcia Z. Wile, Ph.D.
School of Medicine
Case Western Reserve University

Session Number 9.03
American Educational Research Association
Annual Meeting, April, 1974

An Evolving Product Evaluation: Formative and Summative Components

Marcia Z. Wile, Ph.D.

There is a need for a critical system of evaluation and modification of educational products. Based on a four-year research project, the Educational Products Information Exchange Institute reported "an estimated 99 per cent of the learning materials now on the market have never been systematically and effectively shaped and reshaped by feedback from learners." (Komoski, 1971). According to Glass (1972), the methodology for evaluating educational products is not well established and the view that instructional materials are products to be developed and evaluated is relatively recent.

Sanders and Cunningham (1973) offer a general two-dimensional framework in which 1) types and techniques of formative evaluation activities are identified and 2) sources of information for these activities are described.

This paper describes a specific learner-centered product evalua-

---

A paper presented at the American Educational Research Association Annual Meeting, April 17, 1974, Chicago, Illinois.

Dr. Wile is assistant Professor, Division of Research in Medical Education, School of Medicine, Case Western Reserve University, Cleveland, Ohio.

This research was supported in part by NIH grant LM00673 from the National Library of Medicine.

tion model for the formative and summative evaluation of educational materials. With support from the National Library of Medicine, this model has been used by the Division of Research in Medical Education of the School of Medicine, Case Western Reserve University in the development and evaluation of instructional materials for the health professions.

The term product refers to an instructional unit in print or non-print format. Product evaluation relates to the process of determining the merit of such instructional materials from a quality control point of view. Formative components refer to the activities of collecting and analyzing evidence concerning the value of the products while they are in the developmental phases. At this stage, information is fed back to the designers of the materials so that revision can take place. With printed materials, this is prepublication evaluation. Summative components are viewed as the evaluative activities conducted on the final version of the product. With printed matters, this is postpublication evaluation, from which revised editions may evolve.

The Model

The model has four components: learner, author, evaluator, and publisher. The learner-author-evaluator-publisher relationship can be visualized as a triangle with arrows on both ends of each segment, and arrows flowing to and from the center (See Figure 1). The learner is at the center, for he is the core person for whom the material is designed. At the apex is the author who is the subject-expert and creator of the product; the evaluator and publisher are at the baseline, supporting and advising, with the

publisher producing the final product and distributing it. It is dynamic, reactive, and interactive, with feedback from all components providing input to regenerate the process.

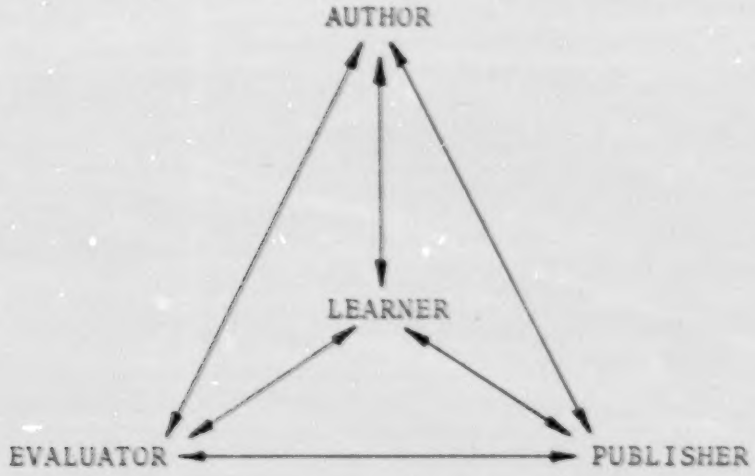

Fig. 1.--The Product Evaluation Model, indicating the components and their relationships.

In this model, the Division of Research in Medical Education, as the evaluator, also serves as the advocate of the learner and the author. The overall goal is to link the expertise and intent of the author with the needs of the learner in the health professions in the creation of instructional materials. The learner-author advocate roles include: identifying and assessing the needs of the learner, eliciting the educational objectives of the units, defining a variety of pedagogic methods, devising prepublication and postpublication evaluation procedures and test instruments, analyzing collected evaluation data for use in the revision of the units, assisting the author in selecting and contracting with a publisher, and seeking publication at the most reasonable price for the learner.

Nine published books and eleven publications in process have been evaluated using this model. In addition, nine units in the formative phases are in the process of prepublication evaluation. Those products which have been evaluated thus far have been printed materials for medical and dental education, However, the model does have applicability to the evaluation of educational materials using other media and emphasizing other disciplines.

The degree of adherence to the model and the types of evaluation approaches are dependent upon the unique characteristics and requirements of the unit. Yet, certain evaluation procedures are common to all educational products in the developmental stages. These are educational review, editorial review, and educational testing.

The educational review examines the content of the unit in relation to its goals, coverage, appropriateness, effectiveness, and learners. This review process attempts to answer such questions as: Why was the unit developed? For whom is the material specifically geared? Who are potential users? Does the material meet the needs of the learners? What is the scope and importance of the material? Is the material accurate and presented with clarity? Are the format, modality, and level of difficulty appropriate for its purposes? Is it liked? Would the users learn from it? Peers, subject experts in intended and related disciplines, intended users, and education specialists participate in these activities.

Editorial review in the prepublication phase focuses upon structure, readability, grammar, and communicative features of the unit. Format and technical aspects of the unit are additional considerations.

Educational testing involves the actual use of the product by students, peers, and learners in different settings. The two types of educational testing approaches used by the Division of Research in Medical Education are field testing and experimental investigations. Field testing includes 1) the informal day-by-day appraisal of the material in actual use by the learner and 2) the more systematic observation and description of the unit by the learner. The systematic experimental studies employ some of the designs categorized by Campbell and Stanley (1963).

Based upon results from the formative evaluation procedures identified above, the materials may be revised, proceed through various evaluation phases, be submitted for publication, or be held in abeyance. Figure 2 illustrates the evaluation cycle.

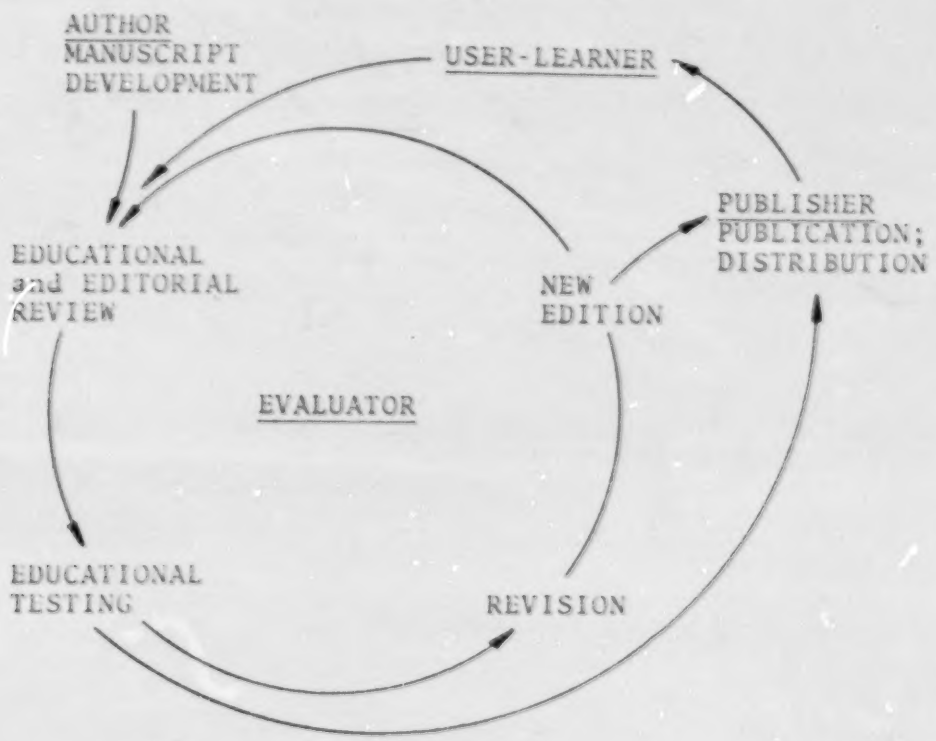

Fig. 2.--The evaluation process cycle showing typical directions of flow.

The next section of the paper describes units representing these approaches to educational product evaluation.

Medical Records, Medical Education and Patient Care

Undoubtedly, the publication which has generated the greatest number of reactions from users and which has been the most responsive to consumer feedback based upon use and review is the book, Medical Records, Medical Education and Patient Care (Weed, 1970). This book presents a system designed to structure the medical record for use in the quality control of medical care and in student-oriented medical education. In 1967, a 141-page lithographed manuscript entitled "The Patient's Record as an Extension of the Basic Science Training of the Physician" was prepared to serve as a preliminary trial of Dr. Weed's hypothesis. One thousand copies of this test-edition were distributed to students, physicians, and individuals in the health professions in hospitals and schools throughout the nation. These users were asked to implement Dr. Weed's techniques and to critique the manual. Based upon appraisals and feedback from these users, the manuscript was revised and produced in a pre-publication paperback edition in 1969. It had been increased to 250 pages of double-spaced text done in offset lithography, and carried the same title as the final version. Five thousand copies of this paperback unit were issued at cost to medical students, faculty, practicing physicians, librarians, and health care professional persons in the United States. The final hardback volume appeared in 1970 and included modifications based upon criticisms and suggestions from the users of both previous editions. Summative or postpublication evaluation

resulted in further refinement. Though no formal instrument was developed to solicit comments, questions and reactions were volunteered from the users who had difficulties and problems in following the Weed system, and were sent to the author and the publisher. In response to this need, Dr. Weed wrote an additional chapter which redescribed certain procedures in simple, direct language. This chapter was incorporated into the fourth printing edition in 1971, and is also available as a separate unit. The Weed book is an outstanding example in which field-testing and reviews from users precipitated revisions from prepublication to postpublication.

The Ocular Fundus in Systemic Disease

Whereas the Weed book was revised four times in as many years, the volume, The Ocular Fundus in Systemic Disease (Chester, 1973) evolved after more than a decade of clinical investigation directed toward the understanding of clinical pathological correlation of retinal abnormalities in systemic disease. The initial material was produced in the form of retinal photographs and photographs of microscopic sections of normal and abnormal retinal tissue. An accompanying syllabus with didactic presentations was introduced during the next five years with the intent of developing an educational self-study unit.

Throughout the developmental phases the unit was reviewed and field-tested by third and fourth year medical students and house officers who were able to apply their learning to actual clinical situations. Feedback from the performance of more than 140 medical students and house officers over a period of ten years resulted in

many modifications. Three major revisions were 1) the addition of explanatory line drawings to clarify the photographs of retinal abnormalities and histologic sections; 2) the development of a series of self-instructional exercises which included case resumes, retinal photographs and/or retinal photomicrographs. These exercises were accompanied by a group of questions and answers which served to test the learner's understanding of previously presented material; and 3) the placement of an annotated bibliography at the end of each section, in addition to the general references.

The revised unit emphasizing self-instruction was used by medical students at all levels, house officers, and practicing physicians. Use of the unit resulted in an improvement of student understanding of the clinical implications of ophthalmic abnormalities. This educational product was subjected to critical peer review by ophthalmologists, pathologists, internists, neurologists, neuropathologists, hematologists, and other subject experts. The evaluation information revealed a favorable response to the unit by learners at all levels, and also that the material, originally designed for medical students and house officers, was of significant value to physicians in many specialties.

As a result of the intense and vigorous evaluation, the unit was submitted for publication. An additional change was made in the format and modality prior to publication. Comments by reviewers indicated that the effectiveness of the slides of the retinal photographs, originally a separate unit, would be enhanced if they were an integral part of the unit. A four-color process was used to transform the color slides to enlarged glossy prints, and the

line drawings were magnified and printed in a striking blue color. These illustrations were made part of the actual text of the published edition. This book is an example of the use of formative evaluation over a period of years to mold an instructional unit.

Two summative evaluation techniques were adopted for this unit. A postage-paid questionnaire, to be returned to the Division of Research in Medical Education, was enclosed in the book. Comments were solicited regarding the use of the book, strong and weak points, and suggestions for change. As of March, 1974, only 25 questionnaires have been returned.

The second evaluation is an investigation now being conducted with a group of first-year medical students. This postpublication evaluation study is an attempt to determine the value of two educational units in the independent learning of basic concepts in ophthalmoscopy, by a sample of students with little experience in clinical medicine, Participation in this study constituted an informal option and was limited to 18 student volunteers. The students completed a pre-test consisting of drawing and labeling structures on the ocular fundus after viewing slides of two normal fundi and were assigned to two groups. One group of nine students received a copy of The Ocular Fundus in Systemic Disease (Chester, 1973); the second group was given a copy of a reference book entitled Textbook of the Fundus of the Eye (Ballantyne and Michaelson, 1970). Each student was given a set of 45 photographic slides illustrating the normal ocular fundus and retinal abnormalities and accompanying study questions. The students were requested to work through the slides, study questions, and the assigned

reference book by themselves. A comparison of the two reference books was not intended. Rather, the focus was on the usefulness of the references in independent learning. A post-test consisting of multiple-choice, true-false, and completion questions will be administered. Selected slides demonstrating unknowns in pathology and clinical abnormalities will constitute a practical test.

An additional instructional unit could result from the summative evaluation study. The guide, composed of questions, answers, illustrations, and a set of slides may be a helpful resource for independent learning. Students' evaluation of the guide will precede further development.

Care of the High-Risk Neonate

Care of the High-Risk Neonate (Klaus and Fanaroff, 1973) discusses major physiologic principles and approaches necessary to the care of the newborn. A unique evaluation feature of this book was the incorporation of evaluation comments from three internationally known subject experts (Doctors Leo Stern, Canada; Samuel Prod'hom, Switzerland; and Gerard B. O'Dell, U.S.A.) into the body of the text. Their critiques were used not only for product evaluation, but also served the purpose of demonstrating the controversial aspects of managing the care of the neonate. Such points of disagreement, attempted clarification, and amplification thus became an integral part of the book and added significantly to its worth.

The 1972 prepublication edition was reviewed by a group of 54 medical and nursing students, staff nurses, house officers, practicing pediatricians, clinical faculty, and specialists in neonatology. The assessment consisted of a questionnaire concerning the content,

readability, self-instructional aspects, and usefulness of each chapter of the book. Open-ended questions sought information about inaccuracies, most important facts learned, and appropriateness of the material.

Overall, the reviewers considered the book more practical than theoretical, worthwhile, and fairly easy to understand with generally complete coverage of specific material. Additionally, the reviewers thought the book was most useful for pediatric house officers followed by pediatricians, practitioners, medical students, and least useful for nurses.

The answers to the open-ended questions were analyzed by the authors and a research assistant. Each of the inaccuracies pointed out by the reviewers was examined by both authors, and corrections were made where appropriate. The evaluation was of assistance in pointing out ambiguities, errors, misstatements, and misinformation. It was interesting that the level of sophistication of the most important information learned reflected the level of education of those reporting. Medical students mentioned basic management procedures and concepts while practicing pediatricians reported more complex and unusual material.

These books illustrate formative evaluation and subsequent modification based upon feedback from field-testing and reviews, and some approaches to summative evaluation. Experimental studies utilize field-testing and reviews, but with the additional component of control.

Drug Abuse and Drug Counseling: A Case Approach

Drug Abuse and Drug Counseling: A Case Approach (Weisman, 1972),

written in 1970 by a medical student, was tested and reviewed in a clinical pharmacology curricular option course by second-year medical students. The one-shot case study technique was used to gather data on performance by these students on a written criterion test. A revised edition was tested with non-randomized experimental and control groups of student counselors working in a state university drug crisis center. Performance on the criterion test of the group which read the workbook was significantly greater than the group which did not read the workbook. In addition to the individuals participating in the studies, the unit was reviewed by physicians, experts in the field, and students. Four revisions occurred before the final published edition. Though the level of control in the studies was not optimal, data were obtained which were valuable in modifying the unit in response to learner needs.

Occlusal Treatment: Preventive and Corrective Occlusal Adjustment

The dental education unit, "Occlusal Treatment: Preventive and Corrective Occlusal Adjustment" written by Sanford C. Frumker, D.D.S. and Norman R. Arnold, D.D.S., is presently in the prepublication evaluation phase. This material, originally designed to teach general practicing dentists the principles of occlusion and techniques of occlusal treatment, has been in preparation for six years.

In November, 1972, the authors approached the Division of Research in Medical Education for assistance with their manuscript which they had developed as a result of their experiences in teaching practicing dentists. Examination of the material resulted in suggestions for pedagogical modifications. Among these recom-

mendations were the inclusion of educational objectives for the entire unit and each chapter, the introduction of summaries and reviews at the end of each chapter, the use of self-evaluation questions and answers, and the development of a glossary, references, and bibliography. It was also suggested that the material be reviewed and tested with dental students as well as with general practitioners.

A prepublication unit was prepared which included objectives, summaries, and self-evaluation exercises. This was sent to subject experts for review. A static-group comparison study was conducted with 32 volunteer senior dental students in the spring of 1973. Sixteen students in the experimental group read the unit, while the sixteen students in the control group had no instruction. Both groups completed a 98-item multiple-choice post-test, assessing concepts and skills. The performance of the students who had read the material was significantly greater ($p < .05$) than the performance of the group who had not, thus indicating that the unit did increase knowledge of occlusal treatment.

Item analysis of the test generated more constructive data than the total test performance information. Questions requiring visualization of the lower jaw moving in relation to the upper jaw were consistently answered incorrectly by both groups of students. Facts alone were insufficient to teach this skill. The learner had to be able to mentally see and feel what was happening in three-dimensional movement. These findings were substantiated in the analysis of the test performance of a group of practicing dentists who had read the unit and had taken the test. They, too, exhibited difficulties in visualizing.

Using the information gathered during the formative evaluation, objectives were rewritten to indicate that the learner, after reading the unit, would know how to do certain procedures, rather than being able to execute these procedures. It was recognized that the learner could not perform the motor skill as a consequence of reading the book, but could be expected to have the information necessary to perform the skill.

Changes were made in the self-evaluation questions and answers and review sections. Visualization was encouraged by the addition of highly descriptive sections which would enable the learner to see structures and movement. For example, in order to promote the visualization of the holding boundary, which is a specific part of the cusp seat, the revised question asked, "Which part of the cusp seat makes sure that your restoration does not allow the mandible to slip forward?" The answer was, "The holding boundary." The original question asked, "What would you do to program the mandibl and give it stability?" The answer was, "Make a holding boundary mesial to the contact point in an upper cusp seat, and distal to the pin-point contact in the lower cusp seat."

Recommendations from reviewers and learners resulted in the inclusion of a bibliography and glossary. Rewriting enhanced the clarity and precision as well as creating a more academic thrust. Evaluation feedback revealed that the unit had potential in an academic setting for the instruction of dental students in addition to its intended use for continuing dental education. A revised test edition has recently been submitted for subject expert review, and negotiations are currently underway with a major publisher.

The true experimental designs are represented by the studies of the use of the self-teaching hematology materials (Harris, et al, 1962; Ginther, 1963) and the evaluation of the book, Human Sexual Behavior: A Workbook in Reproductive Biology (Stenchever, 1970).

Human Sexual Behavior: A Workbook in Reproductive Biology

This semi-programmed workbook with didactic presentation, self-study questions, case studies, and discussions dealing with aspects of human sexuality was written originally as a teaching aid in reproductive biology for medical students. In spring, 1969, a pilot study was designed to explore the usefulness of the unit in increasing knowledge about human sexuality according to performance on a specially constructed 50-item true-false criterion test, "The Assessment of Sex Knowledge." Eighty-eight first-year medical students were randomly assigned to control and experimental groups in a post-test only control group research design. The experimental group members were directed to read the workbook independently, answer the questions, and respond to the case problems. The control group had no task assignment. Both groups were administered "The Assessment of Sex Knowledge." Analysis of data revealed the scores of the experimental group to be significantly greater ($p < .001$) than those of the control group.

The students in the experimental group were asked to complete the "Workbook Assessment Scale," a rating instrument consisting of twenty-eight adjective or word-phrase pairs arranged on a five-point continuum. It was felt that persons having counseling and educational roles in other professions could be potential consumers. This was reinforced by the ratings of the medical students who

indicated the workbook had merit as a learning resource for individuals in the fields of law, education, social work, and medicine.

In addition to the participants in the experimental study, the workbook was extensively critiqued throughout its stages of preparation by faculty, students, and subject experts in the fields of medicine, law, sociology, education, psychology, and theology. Comments and criticisms from these users resulted in considerable modification of the unit.

In fall, 1969, an expanded study which replicated the procedures, design, and materials of the pilot study, was conducted with four professional groups. Seventy post-baccalaureate education students, 126 first-year law students, 95 first-year medical students, and 56 first-year social work students were randomly assigned to experimental and control groups. Analysis of "The Assessment of Sex Knowledge" scores of the control and experimental groups indicated the students who had read the workbook had significantly higher scores ($p < .0001$) than those who had not read the unit.

Additional testing of the workbook and criterion instrument was conducted after the completion of the experimental investigation. Approximately 900 persons participated in the formative evaluation of this instructional unit and in the development of the criterion test.

Other Units

These educational products described have all been published or are in the negotiation for publication stage. However, there have been units which have not resulted in final editions. Sometimes evaluation and testing have indicated that revisions of con-

cepts are required, or that the unit is not suitable for the audience for whom it was intended. A unit on endocrinology, locally developed, was found to have limited application in other institutions with different curricular emphasis. A programmed, problem-oriented printed unit using case problems to teach diagnosis and management was reviewed by two groups of medical students in another medical school. These students generally felt that the print format was unwieldy, that the material presented was available in primary resources and that they would not purchase the unit if given the option. The material is now being evaluated in institutions with computer capabilities for possible use as a computer assisted instructional unit.

Conclusion

Evaluation prior to publication represents an approach to quality control of educational materials. The examples presented in this paper have focused upon the learner centered model. Each instructional unit described has had unique qualities which lent themselves to particular procedures. Whatever the approach, the emphasis has been on the use of evaluation to modify the product in response to learner needs, expectations, experiences, and feedback. This process of input and feedback, and action and reaction among learners, author, evaluator, and publisher has as its ultimate goal the creation of a better instructional product.

References

1. Ballantyne, Arthur J., and Michaelson, Isaac. Textbook of the Fundus of the Eye. Baltimore, Md.: The Williams and Wilkins Company, 1970.

2. Campbell, Donald T., and Stanley, Julian C. Experimental and Quasi-Experimental Designs for Research on Teaching, 171-246, in Gage, N.L., editor, Handbook of Research on Teaching. Chicago: Rand McNally and Company, 1963.

3. Chester, Edward M. The Ocular Fundus in Systemic Disease. Cleveland: The Press of Case Western Reserve University, 1973.

4. Ginther, John R. Cooperative Research in Medical Education: An Example from Hematology. J. Med. Educ., 38: 718-724, 1963.

5. Glass, Gene V. Educational Product Evaluation: A Prototype Format Applied. Educational Researcher, 1: 7, 1972.

6. Harris, John W., Horrigan, Daniel L., Ginther, John R., and Ham, Thomas Hale. Pilot Study in Teaching Hematology with Emphasis on Self-Education by the Students. J. Med. Educ., 37: 719-736, 1962.

7. Klaus, Marshall H., and Fanaroff, Avroy A. Care of the High-Risk Neonate. Philadelphia: W.B. Saunders, 1973.

8. Komoski, P. Kenneth. 50,000,000 Educational Consumers Can't Be Wrong--But Who's Listening. Audiovisual Instruction, 16: 14, 1971.

9. Sanders, James R., and Cunningham, Donald J. A Structure for Formative Evaluation in Product Development. Review of Educational Research, 43: 217-236, 1973.

10. Stenchever, Morton A. Human Sexual Behavior: A Workbook in Reproductive Biology. Cleveland: The Press of Case Western Reserve University, 1970.

11. Weed, Lawrence L. Medical Records, Medical Education, and Patient Care. Cleveland: The Press of Case Western Reserve University, 1970.

12. Weisman, Thomas. Drug Abuse and Drug Counseling: A Case Approach. Cleveland: The Press of Case Western Reserve University, 1972.